Title: Performance of cardiovascular disease risk scores in people diagnosed with type 2 diabetes: external validation using data from the national Scottish diabetes register.

Running title: Diabetes and cardiovascular disease risk

Authors: Stephanie H. Read, PhD[1], Merel van Diepen, PhD[2], Helen M. Colhoun, FRCP[3], Nynke Halbesma, PhD[1], Robert S. Lindsay, FRCP[4], John A. McKnight, FRCP[5], David A. McAllister, MD[6], Ewan R. Pearson, PhD [7], John R. Petrie, MD[4], Sam Philip, MD[8], Naveed Sattar, FMedSci[4], Mark Woodward, PhD[9-11], Sarah H. Wild, FRCP[1], on behalf of the Scottish Diabetes Research Network Epidemiology Group

[1] Usher Institute of Population Health Sciences & Informatics, University of Edinburgh. [2] Department of Clinical Epidemiology, Leiden University Medical Centre. [3] Institute of Genetics and Molecular Medicine, University of Edinburgh. [4] Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK. [5] Metabolic Unit, Western General Hospital, Edinburgh, UK. [6] Institute of Health and Wellbeing, University of Glasgow. [7] Division of Cardiovascular and Diabetes Medicine, University of Dundee, Dundee, UK. [8] Diabetes Research Unit, NHS Grampian, Aberdeen, United Kingdom. [9] The George Institute for Global Health, University of Oxford. [10] The George Institute for Global Health, University of New South Wales, Australia. [11] Department of Epidemiology, Johns Hopkins University, USA.

Correspondence to: Dr Stephanie H Read – Usher Institute of Population Health Sciences & Informatics, Teviot Place, Edinburgh, EH8 9AG (Stephanie.read@ed.ac.uk) TEL: +44 (0)131 6511398, FAX: +44 (0)131 6506868

Word count: 3976

Tables: 3

Figures: 1

Abstract:

Objectives: To evaluate the performance of five CVD risk scores developed in diabetes populations and compare their performance to QRISK2.

Research Design and Methods: A cohort of people diagnosed with type 2 diabetes between 2004 and 2013 was identified from the Scottish national diabetes register. CVD events were identified using linked hospital and death records. Five-year risk of CVD was estimated using each of QRISK2, ADVANCE, Cardiovascular Healthy Study (CHS), New Zealand Diabetes Cohort Study (DCS), Fremantle, and the Swedish National Diabetes Register (NDR) risk scores. Discrimination and calibration was assessed using Harrell's C-statistic and calibration plots, respectively.

Results: The external validation cohort consisted of 181,399 people with type 2 diabetes and no history of CVD. There were 14,081 incident CVD events within five years follow-up. The five-year observed risk of CVD was 9·7% (95% CI: 9·6, 9·9). C-statistics varied between 0·66 and 0·67 for all risk scores. QRISK2 overestimated risk, classifying 87% to be at high risk for developing CVD within five years; ADVANCE underestimated risk and the Swedish NDR risk score calibrated well to observed risk.

Conclusions: None of the risk scores performed well among people with newly diagnosed type 2 diabetes. Using these risk scores to predict five-year CVD risk in this population may not be appropriate.

Introduction:

Despite improvements through earlier diagnoses and improved treatments,[1] cardiovascular disease (CVD) mortality and morbidity risk among people with type 2 diabetes remains markedly higher than in people without diabetes[2, 3]. The effect size depends on the sub-type of CVD as well as age, sex, diabetes duration, ethnicity and socio-economic status.[4, 5]

Accurate CVD risk estimation in people with type 2 diabetes without established CVD can identify patients at high risk of developing CVD and can thus be used to guide appropriate treatment, for example with statins, illustrate to patients the likely effects of lifestyle choices and identify eligible participants for clinical trials. The United Kingdom (UK) clinical guideline network, the National Institute of Health and Clinical Excellence (NICE) recently updated its guidelines to advocate using the QRISK2 score,[6] a risk score developed in the general population,[7] to ascertain CVD risk in people with type 2 diabetes. Despite this recommendation, the performance of QRISK2 has not been independently, externally validated in people with type 2 diabetes.

Several CVD risk scores have also been developed specifically for use among people with type 2 diabetes.[8] While most of the earliest diabetes-specific CVD risk scores, such as the United Kingdom Prospective Diabetes Study (UKPDS) risk engine have been extensively externally validated, many of the contemporary risk scores have not.[8-10] Though one recent study did externally validate several contemporary risk scores,[11] this study was limited by small sample sizes of the external validation cohorts resulting in imprecise estimates of calibration and discrimination.[12, 13] In addition, few external validation studies have been conducted on statin-naïve participants.

Scotland maintains a national register of all patients with a diagnosis of type 2 diabetes, and this register can be linked to population-based hospitalisation and mortality

records. Consequently, this data source offers an opportunity to explore the performance of existing risk scores in a contemporary population of people with type 2 diabetes.

We evaluated the predictive performance of five diabetes-specific CVD risk scores in an external validation cohort of people with type 2 diabetes in Scotland and compared their performance to QRISK2.

Research Design & Methods:

Study design and participants:

Data for these analyses were obtained from the population-wide Scottish Care Information-Diabetes (SCI-Diabetes) database. This dynamic clinical register was established in 2000 and is populated by patient data from primary care and hospital diabetes clinics. Outcome data were obtained from linkage to the Scottish Morbidity records (SMR01), a national hospital admission dataset, and death registrations. Approval for generation and analysis of the linked dataset was obtained from the Caldicott guardians of all Health Boards in Scotland, the Privacy Advisory Committee of the Information Services Division of NHS National Services Scotland (ISD) and the multi-centre research ethics committee.

The external validation cohort consisted of people diagnosed with type 2 diabetes between 1st January 2004 and 1st June 2016 in Scotland. This time-frame was chosen since SCI-Diabetes achieved over 99% completeness of primary and secondary care clinics from 2004 onwards. The cohort was restricted to people who had no previous history of CVD (as defined below) and who were aged between 30 and 89 years at date of diagnosis of diabetes due to small numbers of people in other age groups. We excluded people with a history of CVD at diagnosis of diabetes from our cohort since all but one of the risk scores we wished to validate were designed to estimate risk of incident CVD. We included individuals who were prescribed statins prior to and following type 2 diabetes diagnosis in the main analyses

but conducted sensitivity analyses in sub-populations restricted to i) people who had not been prescribed statins prior to type 2 diabetes diagnosis ii) people who had not been prescribed statins prior to type 2 diabetes diagnosis or during follow-up.

Members of the cohort were followed up from baseline, defined as date of diabetes diagnosis, until date of death, date of first CVD event or study end-date (1st June 2016), whichever came first.

Outcome:

CVD was defined as any hospital admission or death from myocardial infarction, stroke, unstable angina, transient ischaemic attack, peripheral vascular disease and coronary, carotid, or major amputation procedures between baseline and 1st June 2016. International Classification of Disease, version 10 codes and Office of Population Censuses and Surveys Classification of Interventions and Procedures version 4 codes used for identifying CVD may be found in the supplementary material.

Selected risk scores:

QRISK2 was developed using data from the QRESEARCH database, is based upon a Cox proportional hazard model and predicts 10-year risk of CVD.[7] A previous systematic review identified twelve CVD risk prediction models designed for use among individuals with type 2 diabetes.[9] Of these, five (the Swedish national diabetes register risk score (NDR) [14], the Action in Diabetes and Vascular disease: preterAx and diamicroN-MR Controlled Evaluation (ADVANCE) CVD risk score,[15] the Fremantle risk score,[16] the New Zealand Diabetes Cohort study risk score (NZ DCS),[17] and the Cardiovascular Health Study (CHS) risk score[18] were chosen as these were developed to predict CVD while the remaining risk scores predict only coronary heart disease or stroke. Since the publication of the systematic review, an additional risk score, the Atherosclerosis Risk in Communities (ARIC) risk score for CVD has been developed.[19] However, this risk score includes several predictors (alcohol

consumption, physical activity) which are not available in SCI-Diabetes and linked data sources and so was not considered in this validation exercise.

The characteristics of QRISK2 and the five diabetes-specific risk scores are presented in Table 1. All five diabetes-specific risk scores were derived from Cox proportional hazards models; three predict five-year risk, while CHS predicts 10-year risk and ADVANCE predicts four-year risk. The five-year baseline hazard for QRISK2 has been published while the five-year baseline hazards were obtained from the study investigators for CHS and were estimated by extrapolation for ADVANCE.

Predictors used in risk models:

Taken together, the selected CVD risk prediction models contain the following predictors: age, sex, diabetes status (type 1/type 2/no diabetes), diabetes duration, ethnicity, Townsend deprivation score, systolic blood pressure, pulse pressure, smoking status, body mass index, total:HDL-cholesterol ratio, HDL-cholesterol, non-HDL cholesterol, glycated haemoglobin, glucose-lowering medications, micro/macro albuminuria, albumin-creatinine ratio, creatinine, family history of CVD, anti-hypertensive medications, lipid-lowering medications, retinopathy, chronic kidney disease, rheumatoid arthritis and atrial fibrillation.

Definitions of predictors in external validation cohort

Baseline predictor values were defined as measurements recorded closest to baseline, no more than 24 months prior to or 12 months after date of diagnosis of diabetes. Any predictor without a measurement within this timeframe was declared missing. Prescriptions of anti-hypertensive and lipid-lowering medications occurring within the three months preceding baseline date were defined using British National Formulary (BNF) codes 2·5 and 2·12, respectively. Chronic kidney disease was defined as a recording of estimated glomerular

filtration rate of <60ml/min/1·73m^2 and/or a hospital admission for chronic kidney disease (ICD-10 codes: N18, I12-13, ICD-9 codes: 585).

Some predictors were not available, or had different definitions compared to the five scores within SCI-Diabetes and linked datasets and, therefore, some proxy predictors were used. Presence of rheumatoid arthritis was defined as patients with any prescription for disease modifying anti-rheumatic drugs, defined with a BNF code of 10·1·3 prior to baseline. Atrial fibrillation was defined as a hospital admission record, including diagnosis codes for atrial fibrillation (ICD10: I48, ICD9: 427·3) or a warfarin prescription in the absence of a hospital record of prior deep vein thrombosis or pulmonary embolism.[20] For area-based deprivation, the contemporary Scottish measure (Scottish index of multiple deprivation, (SIMD))[21] was mapped across to the historical Townsend score (see Supplementary table 1). Family history was estimated as the conditional probability of having a family history of CVD based on age and deprivation status (SIMD) using data from the 2014 Scottish Health Survey (see Supplementary table 2).[22]

We conducted sensitivity analyses whereby all proxy categorical predictors (atrial fibrillation, rheumatoid arthritis, family history of CVD) were set to null and where the Townsend score was set to the mean. Further sensitivity analyses were conducted to include prevalent diabetes whereby baseline was defined as the latest of 01-01-2010, date of diabetes diagnosis or date of 30th birthday. Lastly, we examined whether the predictive performance of the selected risk scores changed over time (based on diabetes diagnosis before or during/after 2011).

Statistical analyses:

Missing predictor data were imputed using multiple imputation assuming data were missing at random (*mice* package in R)[23]. The imputation model included all predictors and the outcome (follow-up time and CVD event) and was used to generate 20 imputed datasets.

Estimates were pooled using Marshall's adaption of Rubin's rules.[24] Complete case analyses were also conducted as additional sensitivity analyses.

Observed five-year risk of CVD was estimated using the Kaplan-Meier estimator. Five-year risk of CVD was estimated at time of type 2 diabetes diagnosis using the five selected CVD risk scores and QRISK2. The predictive performance of the selected risk scores was assessed by examining measures of calibration and discrimination. Calibration describes how closely the predicted five-year risk and the observed five-year risk agree and was assessed by plotting smoothed observed incidence by predicted incidence using Kaplan-Meier estimates.[25] Calibration-in-the-large statistics and calibration slopes for which values of 0 and 1, respectively, indicate good calibration were also calculated. Calibration-in-the-large statistics compare the mean predicted risk and mean observed risks. Calibration statistics were also calculated for the recalibrated risk scores following adjustment of the baseline hazard to that of the external validation cohort.[26, 27] Discrimination describes the model's ability to differentiate between patients who developed CVD to those that did not and was assessed here by calculating Harrell's C-statistic. This statistic describes the probability that, for any pair of individuals among whom one developed CVD and the other did not develop CVD, the predicted risk of the outcome is higher for the individual who did subsequently develop the disease.[28] A C-statistic of 1 denotes perfect discrimination and a value of 0.5 denotes a prediction model that performs no better than a flip of a coin.

We calculated the number of people classified as high risk, based on the cut-off point in national clinical guidelines (≥10% estimated risk in QRISK2) or low risk (<10% estimated risk in QRISK2).[6]

All statistical analyses were carried out in R version 3·2·2 (2015-08-14). Calibration plots were generated using the *rms* package in R.[29] The reporting of this external validation study

is in accordance with the Transparent Reporting of a multivariable prediction model for individual Prognosis or diagnosis (TRIPOD) guidelines.[30]

Results

There were 218,607 individuals diagnosed with type 2 diabetes in Scotland between January 2004 and June 2016 (Table 2). Of these, 37,208 had a previous history of cardiovascular disease and were excluded from the analyses, leaving 181,399 individuals to form the external validation cohort. Of the 26 predictors included in the risk models, 11 had missing values and the average missingness was 18%. There were a total of 118,098 individuals with incomplete predictor data, including 33,210 individuals with a single incomplete predictor and a further 42,834 individuals with two incomplete variables only (Supplementary Table 3).

Overall, there were 14,081 incident CVD events during 673,740 person-years of follow-up and the five-year observed Kaplan-Meier risk of CVD was 9·7% (95% CI: 9·6, 9·9). The median follow-up time was 5 years and there were 91,549 individuals who were followed-up for at least five years. There were 10,023 non-CVD deaths during follow-up.

Within the external validation cohort, 36,471 individuals had been prescribed statins prior to date of diabetes diagnosis. During follow-up, 71,585 individuals were prescribed statins and the median time until statin initiation was 141 days.

Calibration & Discrimination

Measures of calibration and discrimination are presented in Table 3 and calibration plots are presented in Figure 1. Briefly, the agreement between observed and predicted risks (calibration-in the-large) was better using the Swedish NDR, CHS and NZ DCS risk scores than for the QRISK2 and ADVANCE risk scores. Overall, QRISK2 overestimated risk while ADVANCE underestimated risk across all risk groups. C-statistics for each of the models ranged between 0.663 (0.658, 0.668) and 0.674 (0.669, 0.679) for the whole population. These values decreased following stratification by age, particularly in older age groups. Supplementary Figure 1 presents the distribution of predicted risks for each risk score.

Risk Classification

With a 10% threshold for high risk of developing CVD, QRISK2 classified 86·8% of the cohort as high risk, capturing 13,633 (96·8%) of the subsequent CVD events. In comparison, 3·2%, 58·8%, 25·8%, 82·6% and 37·3% of the cohort were classified as high risk, capturing 8·4%, 80·8%, 59·2%, 94·7% and 46% of the CVD events using the ADVANCE, CHS, Fremantle, NZ DCS and Swedish NDR risk scores, respectively (Supplementary Table 4)

Sensitivity Analyses

After recalibration of the risk scores, calibration improved slightly for the ADVANCE risk score (Supplementary Figure 2). The agreement between observed and predicted risks estimated by QRISK2 deteriorated further. The median predicted risk estimated by the recalibrated QRISK2, ADVANCE, CHS, Fremantle Diabetes Study, NZ DCS and the Swedish NDR risk scores were 94·7%, 2·5%, 4·7%, 4·1%, 6·5% and 6·4%.

Among the subset of individuals who were not prescribed statins prior to diabetes diagnosis (n=144,928), there were 9,572 events during 533,006 person-years of follow-up. Measures of calibration and discrimination for this subset yielded similar results to the main analyses (Supplementary Figure 3 and Supplementary Table 5). These findings were also replicated in the subset of individuals who were not prescribed statins prior to diabetes diagnosis or during follow-up (Supplementary Figure 4, Supplementary Table 5), when proxy variables were replaced with null or mean values (Supplementary Figure 5, Supplementary Table 5), when people with prevalent diabetes were included in the cohort (Supplementary Figure 6, Supplementary Table 6) and when complete case analyses were used omitting missing data (Supplementary Figure 7, Supplementary Table 5). The predictive performance of each of the risk scores varied only slightly depending on year of diabetes diagnosis (<2011 vs. ≥2011) (Supplementary Table 5).

Conclusions:

Using a population-wide diabetes dataset, we have conducted the largest external validation of several contemporary CVD risk scores among people with type 2 diabetes to date and conducted the first external evaluation of QRISK2, the recommended CVD risk score for people with type 2 diabetes in England and Wales.

The ability of the assessed risk scores to discriminate between people who did and did not develop incident CVD as assessed by Harrell's C-statistics was similar with all C-statistics for all risk scores below 0·68. The median predicted risk using QRISK2 was 23·5% compared to an observed risk of 9·3% and QRISK2 classified over 86% of people with type 2 diabetes as high risk. Compared to QRISK2, the agreement between predicted and observed risks using the risk scores developed in diabetes populations were generally better. For example, the median predicted risk using the CHS and Swedish NDR risk scores was 11·7% and 8·3%, respectively. The ADVANCE risk score exhibited the poorest calibration and

severely underestimated risk of CVD in people with type 2 diabetes in Scotland. Recalibration by adjustment of the baseline hazard worsened the performance of QRISK2 since the five-year baseline hazard of the external validation study was higher than the 5-year baseline hazard in the QRESEARCH development dataset. More advanced recalibration approaches in which regression coefficients of the predictors are adjusted, are required to ensure better agreement between QRISK2 predicted and observed risks in people with type 2 diabetes in Scotland.[6] The poor performance of QRISK2 among people with type 2 diabetes could lead to the over-treatment of low risk people.

Findings from other studies:

Although UK national clinical guidelines recommend the use of QRISK2 to estimate CVD risk in people with type 2 diabetes, the performance of QRISK2 in estimating CVD risk in external populations has not previously been assessed. However, an evaluation of the performance of QRISK2 in people with type 2 diabetes has been made using a subset of people with type 2 diabetes in the QRESEARCH database and is described in an online report.[31] This approach to validation, whereby the performance of the model was assessed in a subset of the derivation cohort is likely to have led to optimistic measures of performance. As expected therefore, the C-statistics describing the discriminative ability of QRISK2 were better in this evaluation than in our validation (C-statistics: 0·703 [0·691, 0·715] in women and 0·696 [0·685, 0·706] in men) while the agreement between predicted and observed risks was also better.

Most previous studies have reported that CVD risk scores developed in general populations underestimate risk in people with type 2 diabetes,[8] so we were surprised to find that QRISK2 overestimated risk in our external validation cohort. However, this difference may be partly explained by the inclusion of prevalent type 2 diabetes patients in the QRISK2

derivation cohort, though sensitivity analyses in which people with prevalent type 2 diabetes were included in the external validation cohort did not markedly improve QRISK2's performance (Supplementary Figure 6). Including diabetes in the risk score as a categorical variable and in an interaction with age as in this risk score and others is unlikely to sufficiently capture the complex relationship between diabetes and CVD, particularly the effect of diabetes duration on CVD risk. Similarly, predicting CVD risk in people type 2 diabetes is likely to be further complicated by the possible presence of type 2 diabetes subtypes with distinct disease trajectories.[32] Identifying whether the incorporation of variables denoting type 2 diabetes subtypes within existing risk scores would improve their performance would be of interest for future research.

Previous validation studies of contemporary diabetes-specific risk scores are limited. One recent external validation study assessed the performance of the five diabetes-specific risk scores in three separate cohorts; the European Prospective Investigation into Cancer and Nutrition (EPIC)-NL, EPIC-Potsdam and the Secondary Manifestations of ARTerial disease (SMART) study.[33] Expected to observed ratios varied between 1·06 (0·81, 1·40) and 1·46 (1·04, 2·05). The risk scores exhibited poor discriminative ability in the three external validation cohorts with C-statistics ranging from 0·54 (0·46, 0·63) for the CHS risk score in EPIC-NL to 0·69 (0·59, 0·79) for the Fremantle risk score in SMART. Within each external validation cohort, the discriminative ability was similar for each risk score, a finding replicated in the present study and a possible reflection of the limitations of Harrell's C-statistic in the presence of extensive censoring.[28] Unfortunately, the wide confidence intervals owing to the small numbers of events in each external validation cohort (52 events in EPIC-NL, 73 in EPIC-Potsdam and 58 in SMART) made interpretation of the performance of these models difficult and prevented the authors identifying the strongest performing risk score. The ADVANCE risk score was externally validated in 1,836 patients enrolled in the

DIABHYCAR clinical trial and exhibited similar discrimination (C-statistic: 0·69 [0·65, 0·72]) as reported here, but it underestimated risk in the DIABHYCAR population.[15]

Beyond differences in the performance of different health systems, there are likely to be a number of explanations for the overall poor performance of the assessed risk scores.[34] One major potential explanation is differences in the distribution of outcomes and predictors (i.e. the case mix) in the external validation cohort compared to the derivation cohorts. Different age distributions are likely to be the most important difference between development and this external validation cohort, as indicated by the age-stratified measures of discrimination and calibration in Table 3. A further factor which may have contributed to poor performance in this cohort are different eligibility criteria. For example, ADVANCE was a trial with strict inclusion criteria that made for a very non-standard population.[35] Definitions of CVD also varied between derivation and validation cohorts. While QRISK2 identifies angina through general practice records, the present study only includes hospital admissions for angina and therefore angina incidence will be underestimated. Other factors which may have contributed to the poor performance of these risk scores was the use of proxies, different time frames of the outcome (10-year development vs. 5-year validation) and potentially differences in patterns of glucose-lowering therapies that may have different effects on CVD risks.

Strengths/Weaknesses:

This study had a number of strengths. By utilising population-based registers we were able to assemble the largest external validation cohort of people with type 2 diabetes to assess and directly compare the performance of several CVD risk scores to date. The large cohort also enabled the assessment of each model's performance in subsets of people based upon statin exposure. The population-based nature of these data also ensured low risk of selection biases influencing our findings and enabled us to present results which are applicable to the entire population of Scotland.

A number of weaknesses of the study should be acknowledged. Firstly, the use of proxy measures for some of the predictor variables may have contributed to the poor performance of the models for which these were required. However, by conducting sensitivity analyses to explore the likely effect of using these proxy measures, we have shown that this limitation is unlikely to have had a large effect on the overall findings of our study. Concerns surrounding the accuracy of the recording of CVD events may be a further limitation of this work. Nonetheless, findings from the West of Scotland Coronary Prevention Study (WOSCOPS) indicated that linkage to hospital admissions registers for acquiring CVD events may be as effective as direct patient contact.[36] Finally, we were unable to validate all existing risk scores for people with type 2 diabetes due to the unavailability of some predictors, though risk scores which include variables that are generally not measured may be difficult to implement in clinical practice. We acknowledge that further research is needed to establish whether diabetes treatment contributes to CVD risk independently of other factors. Such research will be particularly valuable for new diabetes drugs that appear to have a beneficial effect on CVD in trial populations.

Implications/Conclusions:

Risk scores have important roles in guiding treatment, communicating risks to patients and for identifying eligible clinical trial participants. Unfortunately, we have shown that many existing risk scores do not accurately predict incident CVD risk in people with newly diagnosed type 2 diabetes, though risk scores developed in diabetes populations generally performed better than QRISK2. Current guidelines which recommend using QRISK2 would classify 87% of people with type 2 diabetes in Scotland as high risk leading to the potential over-treatment of low risk individuals. This approach is therefore not dissimilar to classifying all people aged over 40 years and with type 2 diabetes as high risk, as recommended in several existing clinical guidelines.[37-39]

We conclude that there is scope to improve risk scores for incident CVD among people with type 2 diabetes and suggest that QRISK2 and the five diabetes-specific risk scores, without recalibration, do not currently meet the standard for application to real-world patients in Scotland.

Acknowledgements:

Disclosure Statement: **HC** reports grants, personal fees and non-financial support from AstraZeneca LP Boehringer Ingelheim, Bayer, Eli Lilly, Novartis Pharmaceuticals, Regeneron, Pfizer Inc, Roche Pharmaceuticals, Sanofi Aventis, Sanofi, Novo Nordisk, outside the submitted work. **RSL** reports personal fees from Novo Nordisk, Eli Lilly, Servier outside the submitted work. **EP** reports personal fees from Lilly, MSD and Novo Nordisk outside the submitted work. **NS** reports grants and personal fees from Boehringer Ingelheim, Janssen, Novo Nordisk, Eli Lilly, Amgen, AstraZeneca, Sanofi outside the submitted work. **MW** reports personal fees from Amgen outside the submitted work. **SHR, NH, JP, DAMcA, JMcK, SP & SW** have no disclosures to report.

Contributors: The study was conceived by SHR, SHW, NH, DAMcA, MW, MvD. Data preparation and statistical analyses were carried out by SHR. SHR wrote the first draft of the paper. All authors contributed to the interpretation of the findings and the paper's critical revision. All authors have approved the final version of the manuscript. SHR is responsible for the integrity of the work.

Transparency Statement:

Stephanie Read affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Role of Funder:

Funding for this project came from the Chief Scientist Office (PDF/15/07). The funding source had no role in the design, execution, analysis or interpretation of this study.

References:

1. Hansen, M.B., M.L. Jensen, and B. Carstensen, *Causes of death among diabetic patients in Denmark.* Diabetologia, 2012. **55**(2): p. 294-302.
2. The Emerging Risk Factors, C., *Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies.* Lancet, 2010. **375**(9733): p. 2215-2222.
3. Jackson, C.A., N.R.V. Jones, J.J. Walker, et al., *Area-based socioeconomic status, type 2 diabetes and cardiovascular mortality in Scotland.* Diabetologia, 2012. **55**(11): p. 2938-2945.
4. Shah, A.D., C. Langenberg, E. Rapsomaniki, et al., *Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1·9 million people.* The Lancet Diabetes & Endocrinology, 2015. **3**(2): p. 105-113.
5. Woodward, M., S.A. Peters, and R.R. Huxley, *Diabetes and the female disadvantage.* Womens Health (Lond), 2015. **11**(6): p. 833-9.
6. National Institute for Health and Care Excellence, *Cardiovascular disease: risk assessment and reduction, including lipid modification*. 2014.
7. Hippisley-Cox, J., C. Coupland, Y. Vinogradova, et al., *Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2*. Vol. 336. 2008. 1475-1482.
8. Chamnan, P., R.K. Simmons, S.J. Sharp, et al., *Cardiovascular risk assessment scores for people with diabetes: a systematic review.* Diabetologia, 2009. **52**(10): p. 2001-2014.
9. van Dieren, S., J.W. Beulens, A.P. Kengne, et al., *Prediction models for the risk of cardiovascular disease in patients with type 2 diabetes: a systematic review.* Heart, 2012. **98**(5): p. 360-9.
10. van Dieren, S., L.M. Peelen, U. Nöthlings, et al., *External validation of the UK Prospective Diabetes Study (UKPDS) risk engine in patients with type 2 diabetes.* Diabetologia, 2011. **54**(2): p. 264-270.
11. van der Leeuw, J., S. van Dieren, J.W.J. Beulens, et al., *The validation of cardiovascular risk scores for patients with type 2 diabetes mellitus.* Heart, 2014.
12. Collins, G.S., E.O. Ogundimu, and D.G. Altman, *Sample size considerations for the external validation of a multivariable prognostic model: a resampling study.* Stat Med, 2016. **35**(2): p. 214-26.
13. Vergouwe, Y., E.W. Steyerberg, M.J. Eijkemans, et al., *Substantial effective sample sizes were required for external validation studies of predictive logistic regression models.* J Clin Epidemiol, 2005. **58**(5): p. 475-83.
14. Zethelius, B., B. Eliasson, K. Eeg-Olofsson, et al., *A new model for 5-year risk of cardiovascular disease in type 2 diabetes, from the Swedish National Diabetes Register (NDR).* Diabetes Research and Clinical Practice, 2011. **93**(2): p. 276-284.
15. Kengne, A.P., A. Patel, M. Marre, et al., *Contemporary model for cardiovascular risk prediction in people with type 2 diabetes.* European Journal of Cardiovascular Prevention & Rehabilitation, 2011. **18**(3): p. 393-398.
16. Davis, W.A., M.W. Knuiman, and T.M. Davis, *An Australian cardiovascular risk equation for type 2 diabetes: the Fremantle Diabetes Study.* Intern Med J, 2010. **40**(4): p. 286-92.

17. Elley, C.R., E. Robinson, T. Kenealy, et al., *Derivation and Validation of a New Cardiovascular Risk Score for People With Type 2 Diabetes: The New Zealand Diabetes Cohort Study.* Diabetes Care, 2010. **33**(6): p. 1347-1352.
18. Mukamal, K.J., J.R. Kizer, L. Djousse, et al., *Prediction and classification of cardiovascular disease risk in older adults with diabetes.* Diabetologia, 2013. **56**(2): p. 275-83.
19. Parrinello, C.M., K. Matsushita, M. Woodward, et al., *Risk prediction of major complications in individuals with diabetes: the Atherosclerosis Risk in Communities Study.* Diabetes Obes Metab, 2016. **18**(9): p. 899-906.
20. Morley, K.I., J. Wallace, S.C. Denaxas, et al., *Defining Disease Phenotypes Using National Linked Electronic Health Records: A Case Study of Atrial Fibrillation.* PLOS ONE, 2014. **9**(11): p. e110900.
21. The Scottish Government, *Scottish Index of Multiple Deprivation 2012: A National Statistics Publication for Scotland*. 2012.
22. Brown, L., Christie, S., Gill, V., Gray, L., Hinchliffe, S., Ilic, N., Lepps, H., Leyland, A.H.,, *The Scottish Health Survey*, in *A National Statistics Publication for Scotland*, T.S. Government, Editor. 2014.
23. van Buuren, S., Groothius-Oudshoorn, K., *mice: Multivariate Imputation by Chained Equations in R*. 2011: Journal of Statistical Software, 45 (3), 1-67. .
24. Marshall, A., D.G. Altman, R.L. Holder, et al., *Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines.* BMC Medical Research Methodology, 2009. **9**(1): p. 57.
25. Steyerberg, E.W., *Clinical Prediction Models - A Practical Approach to Development, Validation, and Updating*. 2009, New York: Springer.
26. Crowson, C.S., E.J. Atkinson, and T.M. Therneau, *Assessing calibration of prognostic risk scores.* Stat Methods Med Res, 2016. **25**(4): p. 1692-706.
27. Moons, K.G., A.P. Kengne, D.E. Grobbee, et al., *Risk prediction models: II. External validation, model updating, and impact assessment.* Heart, 2012. **98**(9): p. 691-8.
28. Uno, H., T. Cai, M.J. Pencina, et al., *On the C-statistics for Evaluating Overall Adequacy of Risk Prediction Procedures with Censored Survival Data.* Statistics in medicine, 2011. **30**(10): p. 1105-1117.
29. Harrell, F.E.J., *rms: Regression Modelling Strategies*. 2017: R package version 5.1-0.
30. Collins, G.S., J.B. Reitsma, D.G. Altman, et al., *Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD StatementThe TRIPOD Statement.* Annals of Internal Medicine, 2015. **162**(1): p. 55-63.
31. Hippisley-Cox, J., C. Coupland, and P. Brindle, *Validation of QRISK2 (2014) in patients with diabetes*. 2014.
32. Ahlqvist, E., P. Storm, A. Käräjämäki, et al., *Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables.* The Lancet Diabetes & Endocrinology, 2018. **6**(5): p. 361-369.
33. van der Leeuw, J., S. van Dieren, J.W.J. Beulens, et al., *The validation of cardiovascular risk scores for patients with type 2 diabetes mellitus.* Heart, 2015. **101**(3): p. 222-229.
34. Riley, R.D., J. Ensor, K.I.E. Snell, et al., *External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges.* BMJ, 2016. **353**.
35. The ADVANCE Collaborative Group, *Intensive Blood Glucose Control and Vascular Outcomes in Patients with Type 2 Diabetes.* New England Journal of Medicine, 2008. **358**(24): p. 2560-2572.
36. The West of Scotland Coronary Prevention Study Group, *Computerised record linkage: Compared with traditional patient follow-up methods in clinical trials and illustrated in a prospective epidemiological study.* Journal of Clinical Epidemiology, 1995. **48**(12): p. 1441-1452.

37. Piepoli, M.F., A.W. Hoes, S. Agewall, et al., *2016 European Guidelines on cardiovascular disease prevention in clinical practiceThe Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR).* European Heart Journal, 2016. **37**(29): p. 2315-2381.
38. American Diabetes Association, *8. Cardiovascular Disease and Risk Management.* Diabetes Care, 2015. **38**(Supplement 1): p. S49-S57.
39. Scottish Intercollegiate Guidelines Network, *Risk estimation and the prevention of cardiovascular disease. A national clinical guideline.* 2007.

Tables

Table 1. Characteristics of QRISK2 and five CVD risk scores for use in people with type 2 diabetes

Name	Population	Cohort Type	Time-frame	Follow-up time, years	Main Outcome	Risk Factors	Internal validation C statistic:
QRISK2 risk score	2.3 million people aged 35-74 years in England and Wales without previous CVD	Electronic health records	1993 to 2008	Mean: 7.3 for women. 6.9 for men	10-year risk of CHD, stroke or transient ischaemic attack (ICD-10: I20, I22-I25, I63-I64). Not peripheral arterial disease.	Age, sex, diabetes status, ethnicity, BMI, total:HDL cholesterol, systolic blood pressure, atrial fibrillation, smoking, treated-hypertension, Townsend social deprivation score, Rheumatoid arthritis, family history of CHD.	Men: 0.792 (0.789, 0.794) Women: 0.817 (0.814, 0.820)
Swedish National Diabetes register risk score[14]	24,288 people aged 30-74 years in Sweden	Register	2002 to 2007	Mean: 4.8	5-year risk of fatal or non-fatal CVD. Non-fatal CHD (ICD10: I20-I21) PCI or CABG, fatal CHD (I20-I25), or non-fatal or fatal stroke (I61, I63, I64).	Age, sex, diabetes duration, BMI, total:HDL cholesterol, SBP, HbA1c, smoking, treated hypertension, lipid-lowering drugs, Micro & macro-albuminuria, previous history of CVD	0.72
ADVANCE CVD risk score [15]	7,168 people aged ≥55 years without previous CVD from 215 collaborating centers in 20 countries	Trial	Recruitment: 2001 to 2003	Mean: 4.5	4-year risk of fatal or non-fatal MI or stroke or cardiovascular death. ICD-9 codes for non-fatal event: 430-435, 437-438, 410. ICD-9 codes for fatal event: 394-459, 798.9.	Age, sex, diabetes duration, HbA1c, atrial fibrillation, treated hypertension, albumin-creatinine ratio, pulse pressure, retinopathy, Non-HDL cholesterol	0.70 (0.68, 0.73)

Name	Population	Cohort Type	Time-frame	Follow-up time, years	Main Outcome	Risk Factors	Internal validation C statistic:
	from Asia, Australia, Europe, and North America.						
Fremantle Diabetes Study risk score [16]	1,240 people with a mean age of 64.1 years from Fremantle, Western Australia	Observational cohort study	Recruitment: 1993 to 1996. Follow-up until 2006	Mean: 4.5	5-year risk of fatal or non-fatal MI, stroke or sudden death (No ICD codes provided)	Age, sex, ethnicity, prior CVD, Glycated haemoglobin, Albumin-creatinine ratio, HDL-cholesterol,	0.80
The New Zealand Diabetes Cohort Study (NZ DCS) risk score [17]	36,127 people with a median age of 59 years and without previous CVD from New Zealand	Observational cohort study	2000 - 2009	Median: 3.9	First fatal or non-fatal CVD event, and coronary and peripheral arterial procedures (See: link)	Age, sex, diabetes duration, ethnicity, total:HDL cholesterol, systolic blood pressure, Glycated haemoglobin, smoking, albuminuria	0.68
Cardiovascular Health Study risk score[18]	782 men people aged over 65 and without previous CVD from four field centres in the United States	Observational cohort study	Recruitment between 1989 and 1993. Follow-up until 1999	Mean: 7	10-year risk of MI, stroke and death (No ICD codes provided)	Age, sex, smoking status, HbA1c, systolic blood pressure, total cholesterol, HDL-cholesterol, creatinine, use of glucose-lowering medications	0.64

CHD: Coronary Heart Disease, BMI: Body mass index, HDL-cholesterol: High density lipoprotein cholesterol, PCI: Percutaneous coronary intervention, CABG: Coronary Artery Bypass Grafting, SBP: Systolic blood pressure, HbA1c: Glycated haemoglobin, MI: Myocardial infarction.

Table 2. Baseline characteristics of individuals diagnosed with type 2 diabetes in Scotland between 2004 and 2016 by subsequent five-year CVD outcome status over a median follow-up of 4.9 years

Characteristic		CVD event	No CVD event
N		14,081	167,318
Median age at diagnosis, yrs (IQR)		66.5 (17.4)	59.3 (18)
Sex (%)	Men	8,292 (8.4)	90,604 (91.6)
	Women	5,789 (7)	76,714 (93)
Ethnicity	White	9,808 (7.6)	118,633 (92.4)
	SE-Asian	220 (5.4)	3,836 (94.6)
	Other	480 (6)	7,579 (94)
SIMD (%)	Most Deprived	3,700 (8.4)	40,349 (91.6)
	2	3,361 (8.2)	37,867 (91.8)
	3	2,780 (7.6)	33,569 (92.4)
	4	2,463 (7.5)	30,576 (92.5)
	Least Deprived	1,777 (6.6)	24,957 (93.4)
Mean systolic blood pressure, mmHg (SD)		139.9 (19.9)	138.6 (17.7)
Mean pulse pressure, mmHg (SD)		60.4 (16.3)	57 (14.7)
Smoking status (%)	Current smoker	3,854 (9.7)	35,946 (90.3)
	Ex-smoker	5,463 (8.9)	56,232 (91.1)
	Never smoker	4,699 (5.9)	74,493 (94.1)
Mean BMI, kg/m (SD)		31.3 (6.5)	32.9 (6.9)
Mean total:HDL cholesterol ratio (SD)		4.5 (1.6)	4.7 (1.6)
Non-HDL cholesterol ratio, mmol/mol (SD)		3.9 (1.3)	4.1 (1.3)
Mean glycated haemoglobin, mmol/L (SD)		64 (23)	64.8 (23.4)
Mean glycated haemoglobin, % (SD)		8.0 (4.1)	8.1 (4.2)
Albuminuria (%)	Normal	5,664 (6.6)	80,735 (93.4)
	Micro	1,466 (9.3)	14,333 (90.7)
	Macro	215 (13.6)	1,361 (86.4)
Albumin-creatinine ratio (SD)		5.3 (18.2)	3.3 (12.4)
Prescribed anti-hypertensive medications (%)	Yes	6,053 (9.3)	58,958 (90.7)
	No	8,028 (6.9)	108,360 (93.1)
Prescribed rheumatoid arthritis medications (%)	Yes	210 (9.7)	1,946 (90.3)
	No	13,871 (7.7)	165,372 (92.3)
Atrial Fibrillation (%)	Yes	1,487 (17.3)	7,098 (82.7)
	No	12,594 (7.3)	160,220 (92.7)

Retinopathy (%)	Yes	1,735 (9.2)	17,068 (90.8)
	No	12,346 (7.6)	150,250 (92.4)
Chronic Kidney Disease (%)	Yes	3,547 (13.6)	22,454 (86.4)
	No	9,712 (6.7)	135,537 (93.3)
Prescribed statins prior to diabetes diagnosis (%)	Yes	4,509 (12.4)	31,962 (87.6)
	No	9,572 (6.6)	135,356 (93.4)

Table 3. Age-stratified calibration and discrimination statistics for QRISK2 and five diabetes-specific risk scores

Risk score	Age group	Observed 5-year risk	Median predicted 5-year risk, % (IQR)	Calibration-in-the-large	Calibration slope	C-statistic (Discrimination)
QRISK2	Overall	9.7	24.07 (21.21)	-0.14	0.376 (0.376 ,0.377)	0.674 (0.669, 0.679)
	30-45	3.4	8.73 (9.71)	-0.06	0.208 (0.208 ,0.208)	0.666 (0.644, 0.689)
	46-60	6.8	18.26 (13.81)	-0.11	0.272 (0.272 ,0.273)	0.632 (0.623, 0.641)
	61-75	11.5	29.51 (16.54)	-0.19	0.317 (0.317 ,0.317)	0.604 (0.597, 0.612)
	>75	21.0	45.01 (17.55)	-0.24	0.374 (0.374 ,0.375)	0.578 (0.568, 0.588)
ADVANCE	Overall	9.7	2.00 (2.53)	0.08	1.808 (1.805 ,1.811)	0.666 (0.661, 0.671)
	30-45	3.4	0.58 (0.45)	0.02	3.283 (3.277 ,3.289)	0.628 (0.605, 0.651)
	46-60	6.8	1.33 (0.96)	0.06	2.353 (2.350 ,2.356)	0.595 (0.586, 0.605)
	61-75	11.5	2.93 (2.09)	0.08	1.657 (1.655 ,1.660)	0.594 (0.587, 0.602)
	>75	21.0	6.27 (4.57)	0.15	0.973 (0.970 ,0.976)	0.575 (0.565, 0.585)
CHS	Overall	9.7	11.71 (11.17)	-0.02	0.631 (0.631 ,0.632)	0.674 (0.669, 0.679)
	30-45	3.4	4.58 (2.92)	-0.02	0.760 (0.759 ,0.760)	0.638 (0.615, 0.661)
	46-60	6.8	8.68 (5.34)	-0.02	0.742 (0.742 ,0.742)	0.622 (0.613, 0.632)
	61-75	11.5	16.1 (9.64)	-0.05	0.546 (0.545 ,0.547)	0.603 (0.596, 0.611)
	>75	21.0	26.17 (15.56)	-0.05	0.398 (0.396 ,0.400)	0.575 (0.565, 0.585)
Fremantle Diabetes Study	Overall	9.7	5.24 (7.63)	0.05	0.738 (0.737 ,0.738)	0.665 (0.660, 0.670)
	30-45	3.4	1.2 (0.88)	0.02	2.025 (2.023 ,2.027)	0.626 (0.603, 0.648)
	46-60	6.8	3.23 (2.2)	0.04	1.157 (1.156 ,1.159)	0.591 (0.582, 0.600)
	61-75	11.5	8.49 (5.52)	0.03	0.736 (0.735 ,0.736)	0.593 (0.585, 0.600)
	>75	21.0	20.63 (12.11)	0.00	0.497 (0.496 ,0.497)	0.580 (0.570, 0.590)

NZ DCS	Overall	9.7	16.17 (10.87)	-0.06	0.725 (0.725 ,0.725)	0.670 (0.665, 0.674)
	30-45	3.4	7.7 (2.9)	-0.05	0.679 (0.676 ,0.683)	0.645 (0.622, 0.667)
	46-60	6.8	12.67 (4.37)	-0.06	0.740 (0.739 ,0.741)	0.609 (0.599, 0.618)
	61-75	11.5	20.23 (6.39)	-0.09	0.725 (0.725 ,0.726)	0.599 (0.591, 0.606)
	>75	21.0	30.45 (8.42)	-0.09	0.635 (0.633 ,0.638)	0.573 (0.563, 0.583)
Swedish NDR	Overall	9.7	8.26 (6.79)	0.02	0.955 (0.954 ,0.955)	0.663 (0.658, 0.668)
	30-45	3.4	3.67 (2.26)	-0.01	0.871 (0.871 ,0.871)	0.632 (0.609, 0.654)
	46-60	6.8	6.44 (3.56)	0.01	0.869 (0.869 ,0.870)	0.602 (0.592, 0.611)
	61-75	11.5	10.54 (5.62)	0.00	0.727 (0.727 ,0.727)	0.589 (0.582, 0.596)
	>75	21.0	16.79 (8.74)	0.04	0.576 (0.575 ,0.576)	0.566 (0.556, 0.575)

Figure Legends:

Figure 1: Calibration plots for observed vs. predicted 5-year risk of CVD as estimated using QRISK2, ADVANCE, CHS, Fremantle Diabetes Study, New Zealand Diabetes Cohort Study and Swedish National Diabetes Register risk scores in people diagnosed with type 2 diabetes between 2004 and 2016 in Scotland†

† Grey dashed line reflects perfect agreement between observed and predicted risk